Facial Expression of Pain: Sex Differences in the Discrimination of Varying Intensities

Marie-Pier Plouffe-Demers[1,2], Camille Saumure[1], Daniel Fiset[1], Stéphanie Cormier[1] & Caroline Blais[1]*

[1] Département de Psychoéducation et de Psychologie, Université du Québec en Outaouais.
[2] Département de Psychologie, Université du Québec à Montréal.

Autor Note

Marie-Pier Plouffe-Demers https://orcid.org/0000-0002-4669-2466
Camille Saumure https://orcid.org/0000-0002-3669-4177
Daniel Fiset https://orcid.org/0000-0003-2111-3269
Stéphanie Cormier https://orcid.org/0000-0002-4278-1314
Caroline Blais https://orcid.org/0000-0001-6055-8485

MPPD, CB, CS and DF conceptualized the study. CB programmed the experimental code in Matlab. MPPD programmed the analysis codes in Matlab. MPPD and CS conducted the data collection, and MPPD conducted all statistical analyses. MPPD and CB wrote the first manuscript draft. All authors interpreted the results and revised the manuscript.

The final dataset and accompanying Matlab code are available on the Open Science Framework, DOI [https://osf.io/9t6eu/?view_only=b2e5d1ea641f4be8a40d1aa9b960513f]. The authors declare no competing interests. We would like to thank Philippe Trudel for his help with data collection. This work was supported by a Canada Research Chair in Cognitive and Social Vision held by Caroline Blais, and by graduate scholarships from the Social Sciences and Humanities Research Council to Marie-Pier Plouffe-Demers and Camille Saumure and by a graduate scholarship from the Fonds de Recherche du Québec - Nature et Technologies to Marie-Pier Plouffe-Demers.

CITATION
Plouffe-Demers, M.-P., Saumure, C., Fiset, D., Cormier, S., & Blais, C. (2022, September 8). Facial Expression of Pain: Sex Differences in the Discrimination of Varying Intensities. Emotion. Advance online publication. http://dx.doi.org/10.1037/emo0001156

Abstract

It has been proposed that women are better than men at recognizing emotions and pain experienced by others. They have also been shown to be more sensitive to variations in pain expressions. The objective of the present study was to explore the perceptual basis of these sexual differences by comparing the visual information used by men and women to discriminate between different intensities of pain facial expressions. Using the data-driven Bubbles method, we were able to corroborate the woman advantage in the discrimination of pain intensities which did not appear to be explained by variations in empathic tendencies. In terms of visual strategies, our results do not indicate any qualitative differences in the facial regions used by men and women. However, they suggest that women rely on larger regions of the face which seems to completely mediate their advantage. This utilization of larger clusters could indicate either that women integrate simultaneously and more efficiently information coming from different areas of the face or that they are more flexible in the utilization of the information present in these clusters. Women would then opt for a more holistic or flexible processing of the facial information, while men would rely on a specific yet rigid integration strategy.

Keywords: Sex Differences, Visual Perception, Facial Expressions, Pain Intensity, Data-driven Methods

Facial Expression of Pain: Sex Differences in the Discrimination of Varying Intensities

Communication of pain has been tied to the evolution of the human race as it increases its chance of survival (Prkachin et al., 1983). Communicating pain may alert observers of potential proximal threat and most importantly, lead to caregiving behavior towards the person suffering (Hadjistavropoulos et al., 2011; Yamada & Decety, 2009). Of the many ways to communicate pain, facial expression is one of the most effective (Craig et al., 2011; Williams, 2002). As for other facial expressions, the interpretation of pain facial cues has been shown to be influenced by the context as well as well as the observer's intrapersonal characteristics (Goubert, Vervoort, & Crombez, 2009; Goubert, et al., 2009; Hadjistavropoulos, et al.,1997).

In the general field of emotions, it has been proposed that women are better than men at recognizing emotions experienced by others (Hall, 1978; Kret, 2012; McClure, 2000; Sasson et al., 2004; Thayer & Johnsen, 2000; Thompson et Voyer 2014; Wingenbach, Ashwin & Brosnan, 2018; see however Grimshaw et al., 2004; Palermo & Colthearth, 2004). Interestingly, this has also been suggested for the recognition of facial expressions of pain (Hill & Craig, 2004; Keogh, 2014; Prkachin, Mass & Mercer, 2004; see however Simon et al. 2006; Riva et al., 2011). In addition, women have been shown to be more sensitive to variations in pain expressions than men, which leads to less underestimation bias (Miron-Shatz et al., 2020; Prkachin, Mass & Mercer 2004; Robinson & Wise, 2003).

Several theoretical frameworks have been developed in attempts to explain this feminine advantage such as evolutionary caretaker theories, neurological/hormonal theories and social learning theories (Brody, 1985; Keogh 2014; Kret 2012). However, few studies have explored the visual strategies adopted by women that could potentially be linked to their greater efficiency in facial emotion recognition (Hall et al., 2010; Vassalo et al., 2009). This gap in the literature is surprising given that many studies have reported visual strategy anomalies in patients presenting facial expression recognition disorders such as social anxiety (Faghel-Soubeyrand et al. 2020; Langner, Becker & Rinck, 2009;), schizophrenia (Clark, Gosselin & Goghari, 2003; Faghel-Soubeyrand et al. 2020; Lee et al., 2011), autism (Jones & Klin, 2013; Newmann et al. 2006; Spezio et al. 2008) and prosopagnosia (Fiset et al., 2017; Richoz et al.,2015). Another difference reported in the literature between men and women involves empathic tendencies. Women tend to obtain a higher score on empathy tests (Baez et al., 2017; Batchelder et al., 2017; Derntl et al., 2010; DiTella 2020; Lucas-Molina et al., 2017; Mestre et al., 2009; Reniers et al., 2011; Rueckert et al., 2011; Van der Graaff et al., 2014), which has been in turn linked to the evaluation of facial expressions of emotion and pain (Besel & Yuille, 2010; Kang, Ham & Wallraven, 2016) and should thus be considered in the equation.

Among the studies measuring sexual differences in face perception, some have shown that infant girls aged less than two days old present a stronger interest for faces, whereas infant boys prefer looking at the picture of a mechanical object (Connellan, 2000). Girls also tend to make significantly more eye contact with their parents than boys, and this behavioral difference has been associated with fetal testosterone level (Lutchmaya et al., 2002). Women have also been shown to make significantly more fixations across face images in a face memory task, and this increased scanning has been associated with a better face recognition performance (Heisz, Pottruff & Shore, 2013). Another study that looked at the temporal dynamics of gaze of adult participants has shown that women tend to explore faces more than men, making shorter fixations and larger saccades (Coutrot et al., 2016). Among the studies that have reported a feminine advantage in facial expression recognition (accuracy and/or speed), some have also shown a women's tendency to fixate more the eyes region (Hall et al., 2010) in comparison to men who seemed to further fixate the nose and mouth area (Vassalo et al., 2009). This sex difference, with respect, to fixated regions was also observed by Sæther et al. (2009) in a sex categorization task. Although, this result is not systematic across studies as some reported no such sex differences (Coutrot et al., 2016; Rogers et al., 2017; Sokhn et al., 2017). To our knowledge only one study looked at the link between gaze strategy and empathic tendency and did not find conclusive results (Hall et al., 2010).

While the association between gaze strategy and ability has long been assumed, the link between the two seems to be more tenuous than previously thought. This might account for the discrepancy revealed in the previously mentioned eye tracking results. Even though visual strategy alterations have often been observed in patients with emotion recognition deficits, studies that have directly examined the link between gaze exploration and emotion recognition ability did not report such evidence (see Yitzhak, Perttzovand & Aviezer, 2021 for a review). It is thus possible that the feminine advantage in emotion and pain recognition stems from a particular visual strategy that could not be captured by an eye tracking paradigm. In fact, it has been shown that the overlap between gaze position and visual information extraction is not perfect (Arizpe et al., 2012; Blais et al., 2017; Jonides & Yantis 1981; Posner, 1980; Tardif et al., 2017). For instance, during face processing, one may fixate the center of a face while processing the peripheral information contained in the eyes and/or mouth areas (Caldara, Zhou & Miellet, 2010; Blais et al., 2017; Peterson & Eckstein, 2012;). For these reasons, the utilization of visual information should be directly measured and evaluated as a potential explanation for women's efficiency in facial expression recognition.

The objective of the present study was to compare the visual information used by men and women to discriminate the intensity of pain facial expressions. For this purpose, the data-driven Bubbles method was employed (Gosselin & Schyns, 2001). We expected women to rely on different visual strategies in order to discriminate between different pain intensities and we anticipated these differences of visual information utilization to explain the feminine advantage previously reported. Based on previous eye tracking results suggesting a greater fixation of the eyes area in women (Hall et al., 2010), we could expect qualitative differences in the regions used by both sexes. By fixating the eyes more often, women might have developed an expertise at using those key facial features. Also, studies proposing a more exploratory scan path of faces by women (Coutrot et al., 2016) could predict quantitative differences in the strategies. Women making shorter fixations and larger saccades might be more efficient at integrating facial information from wider regions of the face. As for empathic tendencies self-report questionnaires were administered (Empathy Quotient, Baron-Cohen & Wheelwright, 2004; Interpersonal Reactivity Index, Davis, 1983). We did not expect empathy to drive these sexual differences in terms of visual strategies. Although results previously mentioned report a connection between sex and empathic tendencies and independently, between empathic tendencies and emotion recognition ability, no clear evidence of a link between participants' empathy scores and gaze strategy has yet been found (Hall et al., 2010). However, the current state of knowledge on this topic remains sparse and because of this, we have decided to control for this variable in the analyses (Besel & Yuille, 2010; Kang, Ham & Wallraven, 2016).

Method

Data Availability

The final dataset and accompanying Matlab code are available on the Open Science Framework, DOI [https://osf.io/9t6eu/?view_only=b2e5d1ea641f4be8a40d1aa9b960513f]

Participants

Seventy-six participants (38 males; 22.8 years old on average; SD = 4.6) took part in this study. The sample size was determined *a priori* in order to have a statistical power of .80, assuming a medium effect size of Cohen's f =.25 (G * Power; Faul et al., 2007). All participants had normal or corrected-to-normal visual acuity and were naïve to the purpose of the experiment. Participants were recruited and tested at the University of Quebec in Outaouais, Canada. Four participants were excluded from the analysis because their data were outliers, leaving a total sample size of 72 participants (37 men; see results section for more details). A portion of the data presented in the present study (30 participants; 15 males) has been used as part of a transcultural study which is not published yet and is still under revision. In that study, the visual strategies of 30 East-Asian participants and 30 Western participants were compared. All the participants, from both present study and transcultural study, also completed another task that was not discussed in this article. The purpose of this task was to measure their mental representation of facial expressions of pain (i.e. Reverse Correlation). However, they went through the same protocol (same number of trials) and completed the tasks in the same order, that is to say, the reverse correlation task first, then the Bubbles task, and finally the empathy questionnaires. In this case, the sample size was not sufficient to verify the culture x sex interaction.

Material and Stimuli

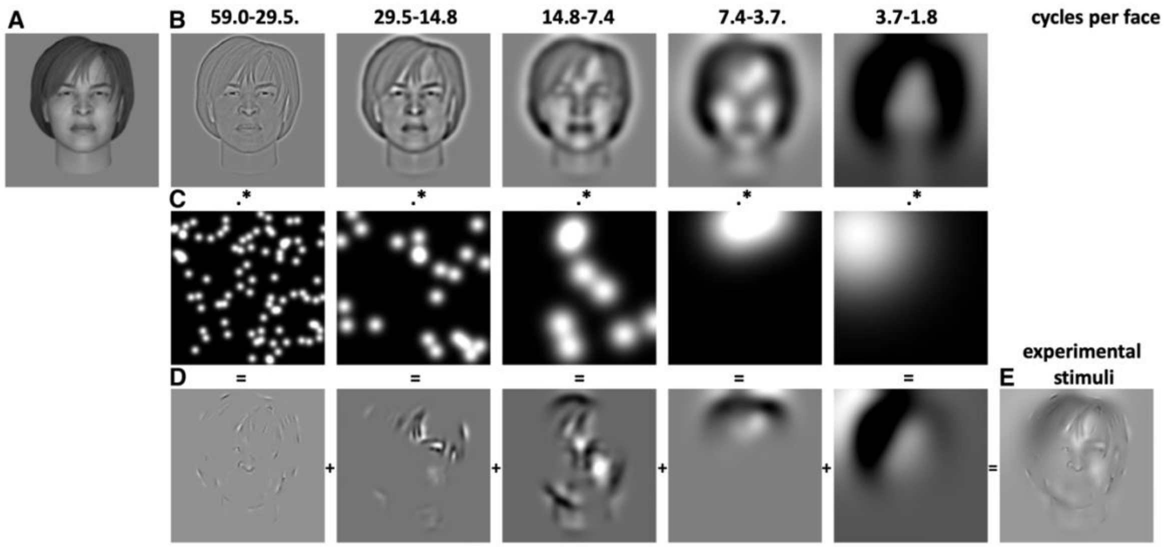
Stimuli were displayed on an LCD (52 × 29 cm; 1920 × 1080p) monitor. All participants were asked to position their head on a chin and forehead rest and all monitors had a calibrated luminance and a refresh rate of 60 Hz. The experimental program was written in Matlab, using functions from the Psychophysics Toolbox (Brainard, 1997; Kleiner, Brainard & Pelli, 2007; Pelli, 1997).

Face stimuli. The faces presented to participants subtended a width of 6° of visual angle (5.3 cm; distance between participants' eyes and screen of 50 cm). Stimuli consisted of avatars created with FACEGen (Singular Inversions Inc., 2009) and FACSGen softwares (Rosech et al., 2011). In total, 8 avatars were created: two genders, in two face ethnicities (White and Asian), each in two emotional states (neutral and the apex of pain facial expression). Those avatars were evaluated by a group of 30 naïve observers who did not take part in the main experiment. Results are consistent with those obtained by Meister et al. (2021), with 100% of ratings from naïve observers being congruent with the FACEGen gender pre-set. As mentioned before, 30 participants (15 males) of this sample took part in another transcultural study, which explains the inclusion of two face ethnicities. However, we were not interested in this variable for the question addressed in the present study. Moreover, analysis on face ethnicity revealed no significant effect on the visual strategy, which allowed us to pool the data across face ethnicities (see results section for more details). In the apex stimuli, three facial action units were activated (i.e. AU 4, 6/7, 9/10). These three features were selected since they have been empirically determined to be frequently found in facial expressions of pain (Kunz & Lautenbacher, 2014; Kunz, Meixner & Lautenbacher, 2019; Prkachin, 1992; Prkachin & Salomon, 2008). We chose to present the most frequent display in order to reduce the number of trials of the experiment even though at least four distinct facial activity patterns of pain resulting from different AU combinations have been proposed (Kunz & Lautenbacher, 2014). For each of the four facial identities, four levels of pain intensity (no pain, 33%, 66%, and 100% of pain) were generated by morphing the neutral and apex avatars (Fantamorph, Abrosoft Co, 2002). As suggested by Krumhuber et al. (2012), these intensity levels can be used in accordance with the 3-point intensity scoring in FACS (low, medium, and high intensity). The final 16 stimuli (2 genders x 2 face ethnicities x 4 levels of pain intensity) were then transformed into grayscale images with a homogeneous gray background. Their luminance was normalized using the SHINE toolbox (Willenbockel et al., 2010).

Bubbles were applied on the stimuli using the following procedure. First, the face picture was decomposed in five frequency bands using the Laplacian pyramid transform implemented in the pyramid toolbox for Matlab (Simoncelli, 1999; 128–64, 64–32, 32–16, 16–8, and 8–4 cycles per image). The entire range of SFs was used, and successive scales were one octave apart, mirroring natural energy statistics and the sensitivity of the human visual system. In this case, the five spatial frequency bands were: 59.0-29.5, 29.5-14.8, 14.8-7.4, 7.4-3.7, and 3.7-1.8 cycles/face, and the remaining low frequency band served as a constant background (see Figure 1 top row for an example of the first step). This step resulted in five images, representing the facial expression stimulus in different spatial resolutions. On each of these five images, bubbles were randomly located; a bubble is a Gaussian aperture through which facial information is made available. The bubbles varied in size as a function of the frequency band, such that their full width at half maximum was of 14.1, 28.3, 56.5, 113.0, and 226.1 pixels from the highest to the lowest spatial frequency band (see Figure 1 middle and bottom rows for an example of the second step). Since the size of the bubbles increased as the spatial scale became coarser, the number of bubbles differed across scales to keep the size of the sampled area constant across frequency bands. Finally, the five sampled images were combined for each face picture to produce the experimental stimuli (see the rightward image of Figure 1 for an example of the experimental stimuli).

Figure 1

Creation of a Stimulus Using the Bubbles Method



Note: Illustration of the procedure to create a stimulus with the Bubbles method. The original face (A) is first decomposed into five spatial frequency bands (B). A mask of randomly positioned Gaussian apertures, called bubbles, is created for each band (C). Each of the five filtered images are then multiplied pixel-by-pixel with their corresponding bubbles mask. The five resulting stimuli (D) are finally fused to create the final stimulus, called bubblized stimulus (E). Thus, in the bubblized stimuli, random facial parts are displayed in different spatial frequencies, allowing to make inference on the facial features and spatial frequencies underlying pain intensity discrimination.

Empathy questionnaires. Participants completed two self-report questionnaires to measure their empathy. Although those two questionnaires are correlated, they do not perfectly overlap (Lawrence et al., 2004), suggesting that they might measure slightly different constructs.

Empathy quotient (EQ). The EQ (Baron-Cohen et al., 2004) is a 60-item questionnaire in which 40 items measure empathy (half of which are reverse-coded), while an additional 20 serve as filler items. Participants responded to each item on a 4-point Likert scale, with higher scores indicating higher agreement. Although the EQ was initially designed to be used in clinical samples, it has also been used in typically developing samples (e.g., Samson 2012; Smith et al. 2010; Krill et al. 2008).

Interpersonal Reactivity Index (IRI). The IRI (Davis, 1980, 1983) is a 28-item questionnaire that was based on the multidimensional theory of empathy (Konrath et al., 2011). It includes four 7-item subscales: empathic concern (EC), perspective taking (PT), fantasy (FS), and personal distress (PD). The EC subscale aims at measuring the other-focused emotional component of empathy. The PT subscale is intended to measure the cognitive component of empathy. The FS items are intended to measure how strongly the respondent identifies with fictional characters in books, movies, or plays. Finally, the PD subscale measures the discomfort and anxiety triggered by others' negative emotional experiences. Participants responded on a scale of 1 to 5, with higher scores indicating higher agreement. IRI, and particularly the EC subscale, has been used in some studies to investigate the potential relationship between pain perception in others and variation in empathy level (Singer et al., 2004; Saarela et al., 2007). IRI also has the advantage of being less explicit than the EQ.

Procedure

After completing a consent and general information form, participants were asked to perform a pain intensity discrimination task which included a total of 3024 trials divided into 21 experimental blocks. They were then asked to fill out the two empathy questionnaires. Since the testing was performed over several hours (three to four hours per participant), most participants needed two experimental sessions (i.e. conducted on two separate days, not necessarily consecutive) to complete all tasks.

At the beginning of each block, instructions were displayed on the screen monitor. Then, on each trial, a fixation cross first appeared in the center of the computer screen for a duration of 500 ms. It was quickly replaced by two bubblized faces of the same identity expressing different pain intensities. Since stimuli were always presented in pairs of the same identity, these variations in intensity enabled the creation of three different levels of difficulty. The easiest level had a 100% difference of intensity between the two stimuli, whereas the hardest had a 33% difference, with an intermediate level of 66% difference. The same bubbles were applied on both stimuli, but the bubbles location varied randomly across trials. Stimuli were displayed on the right and left side of the screen center and remained visible until participants responded. The participants were asked to identify which of the two faces expressed the most pain by pressing the corresponding keyboard key. The key press triggered the next trial. No feedback was provided. The three levels of difficulty, the two genders, and the two face ethnicities occurred an equal number of times within each block, but the order in which they occurred varied randomly from subject to subject. The number of bubbles was adjusted online with an adaptive algorithm (QUEST; Watson & Pelli, 1983) to maintain a target accuracy of 75%. The number of bubbles was adjusted separately for each level of difficulty (33%, 66% or 100% of difference) but was constant across face genders and ethnicities. This adjustment was made to avoid a ceiling effect at the easiest level or a floor effect at the most difficult condition.

The protocol of this experiment was approved by the Research Ethics Committee of Université du Québec en Outaouais and was conducted in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki). All participants provided informed written consent.

Results

Note that in case of sphericity violation the degrees of freedom were adjusted using Greenhouse-Geisser estimate (1959; $\varepsilon \leqq 0.75$) and the Huynh-Feldt estimate (1976; $\varepsilon > 0.75$).

Sex differences in empathic tendencies

Two subscales of the IRI presented significant sex differences and a medium effect size, the FS and the PD (see Table 1 for details). A marginal trend was found with the EC scale of the IRI, which didn't resist the Bonferroni correction (*p* must be <0.01). No significative sex differences were found for EQ scale or other IRI subscales. To ensure that further sex differences were not driven by empathic differences, FS and PD scores were included as covariates in subsequent analysis.

Table 1

Means, Standard Deviations, Independent Sample t Test Significance and Mean Difference Confidence Intervals of EQ and IRI (subscales) Scores for Men and Women

	Sex	*M*	*SD*	*t*	*95%CI [LL, UL]*	*d*
EQ	Men	42.11	12.96	-1.55	[-10.15, 1.28]	-0.37
	Women	46.54	11.23			
EC	Men	35.95	6.61	-1.96	[-6.00, 0.60]	-0.46
	Women	38.91	6.26			
FS	Men	29.11	9.68	-2.73**	[-10.40, -1.60]	-0.64
	Women	35.11	8.97			
PT	Men	35.08	5.83	-1.38	[-5.04, 0.92]	-0.32
	Women	37.14	6.81			
PD	Men	21.24	7.73	-3.16**	[-8.73, -1.98]	-0.74
	Women	26.60	6.61			

Note. EQ, Empathy quotient; EC, empathic concern; FS, fantasy; PT, perspective taking and PD, personal distress. *M* and *SD* are used to represent mean and standard deviation, respectively. Values in square brackets indicate the 95% confidence interval of mean differences. LL and UL stand for confidence interval lower limit and upper limit, respectively. ** indicates $p < .01$.

Ability at discriminating facial expressions of pain as function of the participant's sex

The number of bubbles was used in order to compare men and women's ability to make a distinction between different pain intensities. Since each bubble acts as a window on the stimulus, the number of bubbles represents the amount of spatial information made available to the participant. This means that the participants who needed fewer bubbles were able to successfully complete the task by relying on less facial information. Throughout the experiment, this number of bubbles varied as a function of the participant's accuracy and can thus be used as an index of ability (Royer et al., 2015; 2018). As mentioned in the Participants section, four participants were excluded from the analysis because their mean bubbles number was considered an outlier, using the interquartile rule (i.e. number of bubbles higher than third quartile plus 1.5 multiplied by interquartile range). A 2 x 3 mixed analysis of variance (ANOVA) on the number of bubbles revealed a main effect of the participant's sex, $F(1, 70) = 4.96$, $p = 0.029$, $\eta_p^2 = 0.07$, and a main effect of the level of difficulty, $F(1.16, 81.04) = 225.47$, $p < 0.001$, $\eta_p^2 = 0.76$. No interaction was found between the participant's sex and the level of difficulty, $F(1.16, 81.04) = 1.24$, $p = 0.28$, $\eta_p^2 = 0.02$, indicating a women advantage across all conditions. Women needed on average less bubbles in all three levels of difficulty ($M_{Difficult} = 61.97$, $SD_{Difficult} = 27.42$; $M_{Medium} = 38.78$, $SD_{Medium} = 19.78$; $M_{Easy} = 33.70$, $SD_{Easy} = 17.29$) than men ($M_{Difficult} = 74.77$, $SD_{Difficult} = 31.50$; $M_{Medium} = 50.52$, $SD_{Medium} = 21,05$; $M_{Easy} = 42.90$, $SD_{Easy} = 18.46$) in order to correctly identify the face presenting the higher intensity of pain.

As mentioned before, two subscales of the IRI (i.e. FS and PD) presented significant sex differences. For this reason, we measured the potential links between these empathy scores and the participants' ability. Both Pearsons' correlations were found not to be significant (FS, $r = 0.044$, $p = 0.72$; PD, $r = -0.048$, $p = 0.69$), therefore, we were not able to confirm a link between participants' empathic tendencies and their ability. However, to ensure that the main effect of sex was not driven by any sexual differences in terms of empathy, we conducted a 2 x 3 mixed analysis of covariance (ANCOVA) on the number of bubbles, while controlling for empathy scores (FS and PD). The ANCOVA revealed the same main effect of the participant's sex, $F(1, 68) = 5.66$, $p = 0.02$, $\eta_p^2 = 0.08$, and level of difficulty, $F(1.15, 78.37) = 4.22$, $p = 0.038$, $\eta_p^2 = 0.07$. Again, no interaction was found between the participant's sex and the level of difficulty, $F(1.15, 78.37) = 2.49$, $p = 0.114$, $\eta_p^2 = 0.04$. These results suggest that women are better than men at discriminating between different levels of pain intensity and that this effect is not driven by their empathic tendencies. Given these results, empathy scores were not included in subsequent analyses.

Finally, we conducted an analysis to evaluate the impact of the stimulus gender profile (feminine vs masculine) on participants' discrimination ability. Results suggest that pain expressions were generally better discriminated in feminine faces than in masculine faces and no interaction was found between the sex of the participant and the gender profile of the stimulus. However, since this experiment included only two specimens for each gender profile, the effect of gender cannot be distinguished from the potential effect of the face identity (see section 1 of Supplementary Material for more details).

Sex Differences in the Visual Information Utilization

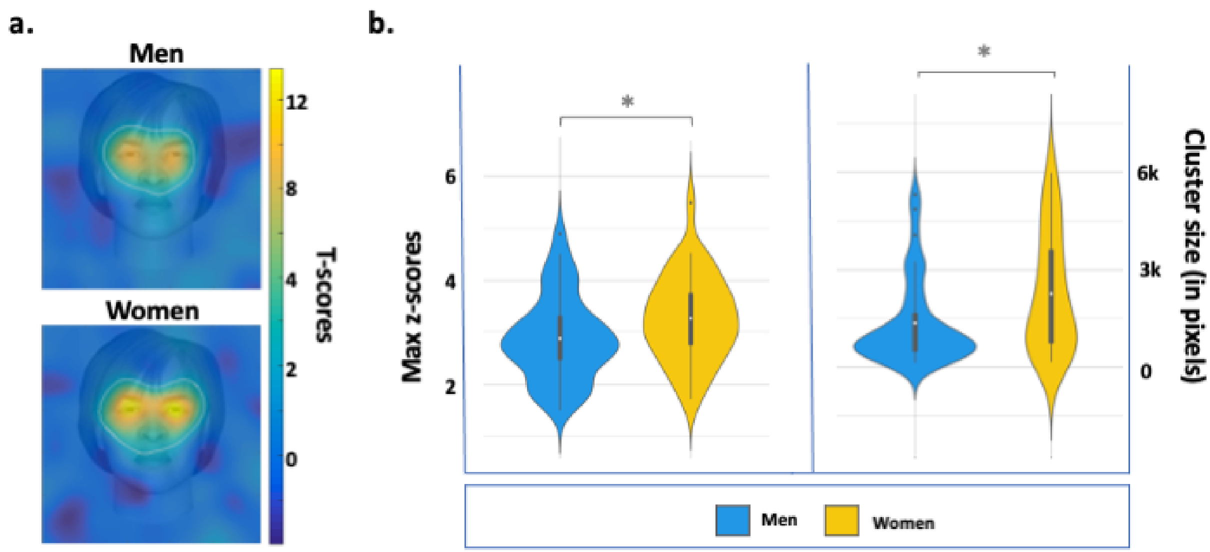
With the Bubbles method, since the bubbles' locations vary randomly across trials, it is possible after a great number of trials to statistically verify the link between the visibility of a pixel (or group of pixels) and the probability that the participant will correctly identify the face displaying the higher pain intensity. To do so, a classification image (CI) is computed for each participant. This CI reveals the visual information (pixels of the image) that was systematically associated with a correct discrimination of pain intensity. By averaging CIs for each sex group, we were then able to measure and compare the visual information used by both groups to discriminate pain intensities. CIs were computed using the following procedure.

First, for each participant, weighted sums of all the bubbles masks used during a given condition of the experiment was calculated. To do so, we used the accuracies transformed into z-scores as weights. This resulted in 30 CI per participants (5 frequency bands x 3 levels of difficulty x 2 face ethnicities) in which facial information increasing the probability of a correct response had a positive value, whereas information decreasing the probability of a correct response had a negative value. In parallel, random CIs were computed using a permutation procedure. This procedure consisted in calculating a weighted sum of all the bubbles masks that were used during a given condition of the experiment, with permuted accuracies transformed into z-scores as weights. These random CIs allowed to estimate the average value and standard deviation expected under the null hypothesis. Secondly, we used those random CIs to transform the participants CIs into Z-scores, for which the Z values indicated the number of standard deviations from chance. Thirdly, the participant's CIs were then averaged across the three conditions and since analyses revealed no significant effect of face ethnicity nor frequency band on the visual strategies (for more details see section 2 and 3 of Supplementary Material), CIs were also averaged across the two face ethnicities and the five frequency bands, resulting in one CI per participant. In this way, by pooling the CIs and thus increasing the number of trials per image, we ensured a better signal-to-noise ratio of the CIs. Finally, CIs were smoothed using Gaussian kernels with a full width at half-maximum of 52 pixels and z-scored again using the averaged and smoothed random CIs.

A pixel-by-pixel bilateral independent-sample t-test was conducted to verify the effect of participants' sex on their utilization of visual information. A Cluster test from the Stat4CI toolbox was applied to control for type 1 error inflation associated with multiple tests (t_{crit} = 2.7, k = 2273.0; p = 0.025; Chauvin et al., 2005). No difference in the regions used by men and women was found [stats, t_{crit} = 2.7, $Cluster_{Max}$ = 251, p = 0.213], both were using mainly the eyes, nose, and upper lips regions (See Figure 2a).

Figure 2

Comparison of the Visual Information Utilization of Men and Women



Note: a) Visual information used by men and women to correctly discriminate between two intensities of pain. Significant areas are delimited by a white contour. The range of colors represent T-score values. Although the background face represents a white female, the analysis was made by combining all trials with no regard to stimulus gender and ethnicity. b) Both sex distributions. The left graph shows the distributions of maximum z-scores. The right graph shows the distribution maximum cluster size in pixels. Men's distributions are depicted in blue and women's in yellow. * indicates $p < .05$

However, an independent-sample t-test on the maximum z-scores suggests that women (M = 3.4, SD = 0.7) tend to be more efficient in the utilization of the information than men (M = 3.0, SD = 0.7), $t(70) = 2.24$, $p = 0.028$, 95% CI [0.04 0.72]; (See Figure 2b left). We also conducted a ROI analysis on the maximum z-scores of the eyes, nose and mouth regions to see if this difference between men and women in terms of efficiency varied across the different parts of the face. The ANOVA did not report any interaction between the participant's sex and ROI, F(1.579, 110.53) = 0.27, p = 0.72, η_p^2 = 0.01. Women were generally more efficient than men for all three regions (for more details see section 4 of Supplementary Material). Moreover, an analysis on the maximum cluster size (in pixels) has shown that women (M = 2262.0, SD = 1337.4) rely on systematically bigger regions than men (M = 1350.0, SD = 1815.20) to discriminate between facial expressions, $t(70)$ = 2.44, p = 0.017, 95% CI [166.41 1659.57]; (See figure 2b right).

Visual Information Utilization and the Ability at Discriminating Facial Expressions of Pain

Pearsons' correlations test revealed that both maximum z-scores and cluster size obtained in each participant's CI were correlated with participant's ability (Cluster size, r = -0.52, p < 0.001; Max z-scores, r = -0.45, p < 0.001). In light of these results, a multiple regression was conducted to investigate whether cluster sizes, maximum z-scores and sex could significantly predict participants' ability. The results of this analysis are displayed in Table 2. The model was significant, F(3,68) = 9.31, p < .001, and explained 29.1% of the variance in ability. While the cluster size contributed significantly to the model (B = -0.42, p = 0.011), the maximum z-scores did not (B = -0.10, p = 0.549). However, a high correlation between the two factors [r = 0.76, p < 0.001] and a VIF of 2.441 raises the possibility of multicollinearity or near dependencies. Surprisingly, the sex of the participant did not significantly contribute to the model either (B = -0.12, p = 0.282).

Table 2

Regression Results Using the Ability as the Criterion

Predictor	*b*	*Beta*	*Beta* 95% CI [LL, UL]	Fit
(Intercept)	77.76**			
Cluster size	-0.01*	-0.42*	[-0.74, -0.10]	
Maximum z-score	-2.91	-0.10	[-0.41, 0.22]	
Sex	-5.20	-0.12	[-0.33, 0.10]	
				R^2 = .291**
				95% CI [.10,.42]

Note. *b* represents unstandardized regression weights. *Beta* indicates the standardized regression weights. *LL* and *UL* indicate the lower and upper limits of a confidence interval, respectively. *r* represents the zero-order correlation. * indicates $p < .05$. ** indicates $p < .01$.

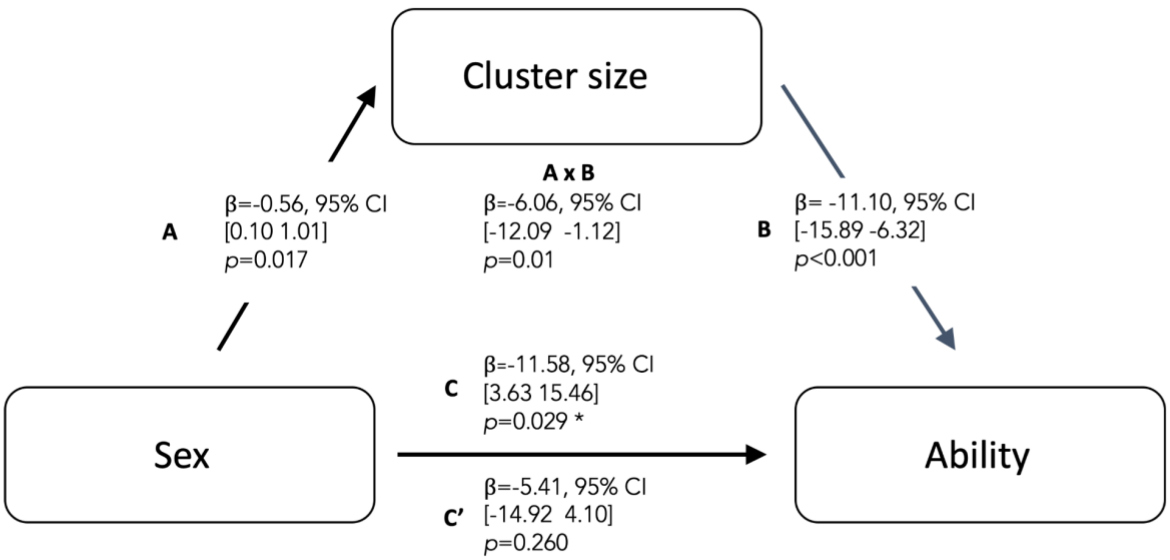
Full Mediation Effect of Information Utilization on Women Advantage

Previous analyses revealed a sex difference in discrimination ability. However, that effect was not significant when other factors (i.e. maximum z-scores and cluster size) were taken into account. One possibility is that the relationship between sex and ability could be explained by their relationship to a third mediator variable. The cluster size was then examined as a mediator of the relation between participants' sex and ability. We conducted a bootstrapped mediation analysis (1000 resamples using "mediation" R Package; Tingley et al., 2013). The indirect

effect of sex (dummy coded: men = 1, women = 0) on ability via cluster size was significant (A x B β = -6.06, 95% CI [-12.09 -1.12] $p = 0.01$; see Figure 3). According to the causal-step approach, individual differences in cluster size significantly and completely mediated the effect of sex on the ability to discriminate pain intensities. To estimate the size of our mediated effect, we computed the effect size of the indirect effect using Lachowicz, Preacher, and Kelley (2018) upsilon estimate and the 95% BCa CIs from 1000 bootstraps using the MBESS R package (Kelley & Lai, 2010). Within Cohen's guidelines (1988), the derived effect size was in the small range (upsilon = 0.019, 95% CI [0.001 0.062]).

Figure 3

Simple Mediation via Cluster Size



Note: A= effect of sex on cluster size. B= effect of cluster size on ability. C = total effect of sex on ability. C'= Direct effect of sex on ability after adding cluster size to the model. AxB = mediation of the effect of sex on ability by cluster size.

We acknowledge that mediation analysis cannot conclusively demonstrate the direction of causality between measured variables. We thus considered an alternative mediation model (Fiedler, Harris & Schott, 2018). In this inverse causal model, the relation between sex and cluster size would be mediated by the amount of information (i.e. number of Bubbles) made available to participants, in which case the reliance on larger clusters would be driven by the method. In the present study the number of bubbles was adjusted in order to maintain a target accuracy of 75%, and since women were found to rely on fewer bubbles to complete the task, it is possible that it led to the utilization of larger clusters. Nevertheless, to verify this alternative model, we conducted a model-observer analysis using the same Bubbles task as the one performed by the participants. For one model-observer, the number of Bubbles was fixed according to the average information utilization of men (44 Bubbles) and for the other, the number of Bubbles was fixed to the average information utilization of women (56 Bubbles). Each model observer completed 10 000 trials. We then proceeded to a permutation analysis. On each of the 1000 permutations, we sampled 3024 bubbles masks from each distribution to generate a CI for each group. We then compared the maximum cluster size used by both model-observers. Results did not reveal any significant difference between the groups (95% CI [-2671, 1341], $p = 0.667$). It is thus not possible to confirm a causal relationship between the number of bubbles made available and the use of larger clusters and it is unlikely that the number of bubbles would then act as an alternative mediator to the relation between sex and cluster size.

Discussion

There is consistent evidence for sex-related effects in the decoding of the facial expression of pain, which, in many cases, suggest the presence of a feminine advantage (Hill et al., 2004; Keogh, 2014; Prkachin et al., 2004;). Among the many frameworks that have been developed in order to explain these differences, few have looked at the visual strategies used by women that could account for their increased ability. In this study, we used Bubbles, a data-driven psychophysical method, to compare the visual information utilization of men and women to discriminate intensities of pain facial expressions. Although sexual differences in terms of gaze pattern (Coutrot et al., 2016; Hall et al., 2010; Heisz et al., 2013; Vassalo et al., 2009) have been previously raised, to our knowledge, this is the first study to directly compare their utilization of facial information.

In terms of ability, our findings corroborate previous results (Miron-Shatz et al., 2020; Prkachin et al., 2004; Robinson et al., 2003) and indicate a woman advantage in the discrimination of pain intensities. More specifically, women need significantly less facial information to complete the task. In terms of visual strategies, our results do not indicate any qualitative differences in the facial regions used by men and women. Both sexes are mainly using the eyes, nose, and upper lip regions to successfully complete the task. However, our results suggest that women rely on larger clusters than men, and that the maximum z-scores of their CIs are systematically higher than those of men's. The presence of multicollinearity between cluster size and maximum z-scores may as well indicate that they both represent the same mechanism, which is most likely a more efficient use of facial information. Therefore, although the results suggest that both factors are related to participants' ability to discriminate pain intensity, the multicollinearity between the two makes it difficult to evaluate their individual contribution to the regression model. Thus, only the cluster size was considered in the subsequent analysis. In addition, the results of the mediation analysis show that the reliance on a larger cluster size completely mediates the advantage found in women in terms of ability. Finally, our results did not indicate that sex differences in the ability to discriminate pain intensities are driven by self-reported empathy scores. This finding is in line with previous studies revealing that higher self-reported empathy is associated with a general tendency to give higher estimations when asked to rate the pain of another but is not necessarily associated with more "accurate" pain ratings (Green et al., 2009).

The reliance on larger clusters could potentially be explained by two mechanisms. Either the participants integrate larger regions (i.e. information coming from different areas of the face) or they are more flexible in the utilization of the information present in that cluster, in which case the presence of any small part of the cluster could lead to a successful discrimination of pain intensities. In this case, smaller clusters found in men CIs might indicate that they rely on smaller but essential regions that must necessarily be revealed for them to successfully complete the task.

As for women, they might rely on larger regions and as long as some parts of the information contained in those regions are present, they are able to correctly identify the face expressing the most intense pain. Men would then opt for a more specific and rigid integration strategy and women would rely on a more holistic or flexible processing of the facial information. The reliance on larger clusters is in line with the aforementioned eye-tracking studies suggesting that women tend to use a more exploratory scanning strategy, make more fixations across face images and present a shorter ratio of fixation duration to saccade duration (Heisz et al., 2013; Coutrot et al.,2016). By actively looking for cues distributed in many parts of the face, women might be able to process and integrate larger areas of the face, hence larger clusters, than men. The flexibility of their information utilization might also allow this exploratory gaze pattern and favor a faster and more efficient inspection of faces.

Among the many frameworks previously proposed, socio-cultural theories could predict a more efficient use of facial information in women. For example, the biosocial constructionist model by Wood and Eagly (2012) suggests that the physiological dimorphism found between men and women created task efficiency differences leading to a division of labor, with women being primarily responsible for child-rearing and men, for gathering and hunting food. In that spirit, women should be more sensitive to nonverbal cues, since nonverbal sensitivity is adaptive to bearing and nursing children. It has also been suggested that parents socialize their sons' and daughters' emotions differently (apply different contingencies to their behaviors) as the norms within a particular culture dictate the masculinity or femininity of specific emotions. For example, some studies from social learning theories have demonstrated differences in the way parents discuss past events with their children, using more emotional words with their daughters than with their sons (Brody, 2000). It is thus possible that women under either socio-cultural or developmental pressures have developed over time a particular interest and expertise for the visual information that seems the most adaptive, like facial expressions of emotion. Another biological hypothesis, also based on the physiological and cerebral dimorphism found between men and women suggest that differences in specific brain regions could potentially explain why men and women tend to perceive the world differently (Vanston & Strother, 2017). For instance, it has been shown that men tend to have larger visual cortex than women (Amunts et al., 2007; Handa and McGivern, 2015). It has also been shown that variability in the primary visual cortical surface (V1) presents a tradeoff between sensitivity to visual details and susceptibility to visual context modulation (Song, Schwarzkopf & Rees, 2013). Individuals with larger V1 tend to discriminate finer orientation differences and tend to be more precise in their mental imagery. In contrast, individuals with smaller V1 are more susceptible to contextual modulations and tend to have a stronger but less precise sensory imagery (Bergmann et al., 2016; Schwarzkopf, Song & Rees, 2011; Song et al., 2013). In light of these results, one possible - although speculative - explanation to the sexual differences found in terms of visual information integration (i.e. clusters size) could rely on these correlations between cortical anatomy and visual perception. As such, women who tend to have a smaller V1 surface would lean toward a more context oriented or holistic processing of the facial expression, while men who tend to have a larger visual cortical area would adopt a more detail-oriented scope. This hypothesis is in line with previous results in which men have been shown to be less context-sensitive in comparison to women (Barnett-Cowan, et al., 2010; Baron-Cohen, 2002; Phillips, Chapman & Berry, 2004; see however Shaqiri et al. 2018).

The results of the present study must be interpreted in light of some limitations. First, hormonal variations in women that could potentially influence their ability to discriminate pain intensities were not controlled for. In fact, pregnancy, the menstrual phase, and the use of oral contraceptives have been shown to impact women's perception of faces, facial expressions, and social cues (see Little, 2013). For example, it has been shown that women in their mid-luteal phase perceive fearful and disgusted expressions as more intense (Conway et al., 2007). Another limitation was the use of artificially generated emotional faces, which could be considered as less ecological. However, this offers the advantage of providing a perfect control over the intensity in which each facial feature (action unit) is activated. In addition, one previous study has directly compared the utilization of different kinds of facial stimuli in a Bubbles paradigm and suggest that the results obtained with avatar faces generalize to real faces (Robinson et al., 2014). Also, the use of avatars has been previously validated in different experimental settings and has been shown to give results similar to those obtained with real faces in pain (Blais et al., 2019; Hirsh, George & Robinson, 2009; Lin et al., 2020; Meister et al., 2021; Riva et al., 2011;Tessier et al., 2019; Wandner et al., 2010;).

It is possible that the use of this type of stimulus, in which action units were varying more systematically, has favored the male strategy, and thus minimized sex differences in terms of ability in comparison to what would be expected in real life. In fact, it has been shown that spontaneous facial expressions are by nature more subtle and ambiguous (Saumure et al., 2018; (Hess & Blairy, 2001; Kayyal & Russell, 2013; Motley & Camden, 1988; Naab & Russell, 2007; Wagner, MacDonald, & Manstead, 1986). For this reason, this experimental paradigm should be tested with real and spontaneous facial expressions of pain. It would also be interesting to test it with dynamic facial expressions since it has been suggested that static and dynamic facial expressions are perceived and processed differently (e.g. eye movements, Blais et al., 2017; spatial frequency utilization, Plouffe-Demers et al., 2019). Also, motion sensitivity has been found to differ between men and women (Vanston et al., 2017). Future research using material featuring human individuals should also consider the impact of the actor's gender profile on sexual differences, since it has been previously demonstrated that the gender of the stimuli could impact the perception of pain (Simon et al., 2006; Simon et al., 2008; Riva et al., 2011). Although most of the results suggest that pain is in general better processed for male faces than for women faces (Coll, et al., 2012; Pronina & Rule, 2014; Simon et al., 2006), our analysis suggests that pain expressions were in this case more accurately discriminated in female-looking faces than in male-looking faces (see section 1 of Supplementary Material for more details). These results should nevertheless be interpreted with caution since our research setting was not initially designed to consider the impact of stimulus gender. Finally, as suggested by Kunz & Lautenbacher (2014), future research should include various displays of pain facial expressions as it is possible that men and women differ in the display they prioritize. Also, since women have also been shown to be more accurate than men in recognizing facial expressions of other emotions than pain (Campbell et al., 2002; Mandal & Palchoudhury, 1985; Montagne, et al., 2005; Whittle et al., 2011), future research should include other facial expressions to verify if the utilization of larger clusters of visual information by women is a finding that can be generalized.

Conclusions

The current study corroborates previous results suggesting a feminine advantage in the processing of pain perceived in others. However, it suggests that the ability in which women were found to better discriminate between different pain intensities do not necessarily rely on the utilization of specific facial features, but rather on a more efficient use of this information. The reliance on larger regions of the face suggests that women are either better at simultaneously integrating information coming from different parts of the face or more flexible in the utilization of this information. This holistic processing of the facial expression would give women an advantage on men who tend to adopt a more detail-oriented scope. Although this study provides a possible perceptive explanation to the female advantage, future research should disentangle these flexibility vs. efficiency hypotheses as well as their potential link with the cortical dimorphism found between men and women.

References

Amunts, K., Armstrong, E., Malikovic, A., Hömke, L., Mohlberg, H., Schleicher, A., & Zilles, K. (2007). Gender-specific left–right asymmetries in human visual cortex. *Journal of Neuroscience*, *27*(6), 1356-1364. https://doi.org/10.1523/JNEUROSCI.4753-06.2007

Arizpe, J., Kravitz, D. J., Yovel, G., & Baker, C. I. (2012). Start position strongly influences fixation patterns during face processing: Difficulties with eye movements as a measure of information use. *PloS one*, *7*(2), e31106. https://doi.org/10.1371/journal.pone.0031106

Baez, S., Flichtentrei, D., Prats, M., Mastandueno, R., García, A. M., Cetkovich, M., & Ibáñez, A. (2017). Men, women… who cares? A population-based study on sex differences and gender roles in empathy and moral cognition. *PloS one*, *12*(6), e0179336. https://doi.org/10.1371/journal.pone.0179336

Barnett-Cowan, M., Dyde, R. T., Thompson, C., & Harris, L. R. (2010). Multisensory determinants of orientation perception: task-specific sex differences. *European Journal of Neuroscience*, *31*(10), 1899-1907. https://doi.org/10.1111/j.1460-9568.2010.07199.x

Baron-Cohen, S. (2002). The extreme male brain theory of autism. *Trends in cognitive sciences*, *6*(6), 248-254. https://doi.org/10.1016/S1364-6613(02)01904-6

Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of autism and developmental disorders*, *34*(2), 163-175. https://doi.org/10.1023/b:jadd.0000022607

Batchelder, L., Brosnan, M., & Ashwin, C. (2017). The development and validation of the empathy components questionnaire (ECQ). *PloS one*, *12*(1), e0169185. https://doi.org/10.1371/journal.pone.0169185

Bergmann, J., Genç, E., Kohler, A., Singer, W., & Pearson, J. (2016). Smaller primary visual cortex is associated with stronger, but less precise mental imagery. *Cerebral cortex*, *26*(9), 3838-3850. https://doi.org/10.1093/cercor/bhv186

Besel, L. D., & Yuille, J. C. (2010). Individual differences in empathy: The role of facial expression recognition. *Personality and individual differences*, *49*(2), 107-112. https://doi.org/10.1016/j.paid.2010.03.013

Blais, C., Fiset, D., Furumoto-Deshaies, H., Kunz, M., Seuss, D., & Cormier, S. (2019). Facial features underlying the decoding of pain expressions. *The Journal of Pain*, *20*(6), 728-738. https://doi.org/10.1016/j.jpain.2019.01.002

Blais, C., Fiset, D., Roy, C., Saumure Régimbald, C., & Gosselin, F. (2017). Eye fixation patterns for categorizing static and dynamic facial expressions. *Emotion*, *17*(7), 1107. http://doi.org/10.1037/emo0000283.supp

Brainard, D. H., & Vision, S. (1997). The psychophysics toolbox. *Spatial vision*, *10*(4), 433-436.

Breedlove, S. M. (1994). Sexual differentiation of the human nervous system. *Annual review of psychology*, *45*(1), 389-418. http://doi.org/10.1146/annurev.ps.45.020194.002133

Brody, L. R. (1985). Gender differences in emotional development: A review of theories and research. *Journal of personality*, *53*(2), 102-149. http://doi.org/10.1111/j.1467-6494.1985.tb00361.x

Brody, L. R. (2000). The socialization of gender differences in emotional expression: Display rules, infant temperament, and differentiation. *Gender and emotion: Social psychological perspectives*, *2*(11), 122-137.

Burt, P. J., & Adelson, E. H. (1983). A multiresolution spline with application to image mosaics. *ACM Transactions on Graphics (TOG)*, *2*(4), 217-236. http://doi.org/10.1145/245.247

Caldara, R., Zhou, X., & Miellet, S. (2010). Putting culture under the 'spotlight'reveals universal information use for face recognition. *PLoS One*, *5*(3), e9708. http://doi.org/10.1371/journal.pone.0009708

Campbell, R., Elgar, K., Kuntsi, J., Akers, R., Terstegge, J., Coleman, M., & Skuse, D. (2002). The classification of 'fear' from faces is associated with face recognition skill in women. *Neuropsychologia*, *40*(6), 575-584. http://doi.org/10.1016/s0028-3932(01)00164-6

Chauvin, A., Worsley, K. J., Schyns, P. G., Arguin, M., & Gosselin, F. (2005). Accurate statistical tests for smooth classification images. *Journal of vision*, *5*(9), 1-1. http://doi.org/10.1167/5.9.1

Clark, C. M., Gosselin, F., & Goghari, V. M. (2013). Aberrant patterns of visual facial information usage in schizophrenia. *Journal of Abnormal Psychology*, 122, 513–519. http://doi.org/10.1037/a0031944

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). New York, NY: Academic Press. http://doi.org/10.2307/2529115

Coll, M. P., Budell, L., Rainville, P., Decety, J., & Jackson, P. L. (2012). The role of gender in the interaction between self-pain and the perception of pain in others. *The Journal of Pain*, *13*(7), 695-703. http://doi.org/10.1016/j.jpain.2012.04.009

Connellan, J., Baron-Cohen, S., Wheelwright, S., Batki, A., & Ahluwalia, J. (2000). Sex differences in human neonatal social perception. *Infant behavior and Development*, *23*(1), 113-118. http://doi.org/10.1016/s0163-6383(00)00032-1

Coutrot, A., Binetti, N., Harrison, C., Mareschal, I., & Johnston, A. (2016). Face exploration dynamics differentiate men and women. *Journal of vision*, *16*(14), 16-16. http://doi.org/10.1167/16.14.16

Conway, C. A., Jones, B. C., Debruine, L. M., Welling, L. L. M., Smith, M. L., Perrett, D. I., ... & Al-Dujaili, E. A. (2007). Salience of emotional displays of danger and contagion in faces is enhanced when progesterone levels are raised. *Hormones and behavior*, *51*(2), 202-206. http://doi.org/10.1016/j.yhbeh.2006.10.002

Craig, K. D., Prkachin, K. M., & Grunau, R. E. (2011). The facial expression of pain. In D. C. Turk & R. Melzack (Eds.), *Handbook of pain assessment* (pp. 117–133). The Guilford Press.

Davis, M.H. (1980). A multidimensional approach to individual differences in empathy. JSAS Catalog of Selected Documents in Psychology, 10, 85.

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of personality and social psychology*, *44*(1), 113. http://doi.org/10.1037/0022-3514.44.1.113

Derntl, B., Finkelmeyer, A., Eickhoff, S., Kellermann, T., Falkenberg, D. I., Schneider, F., & Habel, U. (2010). Multidimensional assessment of empathic abilities: neural correlates and gender differences. *Psychoneuroendocrinology*, *35*(1), 67-82. http://doi.org/10.1016/j.psyneuen.2009.10.006

Di Tella, M., Miti, F., Ardito, R. B., & Adenzato, M. (2020). Social cognition and sex: Are men and women really different?. *Personality and Individual Differences*, *162*, 110045. http://doi.org/10.1016/j.paid.2020.110045

Faghel-Soubeyrand, S., Lecomte, T., Bravo, M. A., Lepage, M., Potvin, S., Abdel-Baki, A., ... & Gosselin, F. (2020). Abnormal visual representations associated with confusion of perceived facial expression in schizophrenia with social anxiety disorder. *NPJ schizophrenia*, *6*(1), 1-9. http://doi.org/10.1038/s41537-020-00116-1

Fiedler, K., Harris, C., & Schott, M. (2018). Unwarranted inferences from statistical mediation tests–An analysis of articles published in 2015. *Journal of Experimental Social Psychology*, *75*, 95-102. http://doi.org/10.1016/j.jesp.2017.11.008

Fiset, D., Blais, C., Royer, J., Richoz, A. R., Dugas, G., & Caldara, R. (2017). Mapping the impairment in decoding static facial expressions of emotion in prosopagnosia. *Social cognitive and affective neuroscience*, *12*(8), 1334-1341. http://doi.org/10.1093/scan/nsx068

Goubert, L., Vervoort, T., Cano, A., & Crombez, G. (2009). Catastrophizing about their children's pain is related to higher parent–child congruency in pain ratings: an experimental investigation. *European Journal of Pain*, *13*(2), 196-201. http://doi.org/10.1016/j.ejpain.2008.03.009

Goubert, L., Vervoort, T., & Crombez, G. (2009). Pain demands attention from others: the approach/avoidance paradox. *Pain, 143(1), 5–6.* http://doi.org/10.1016/j.pain.2009.02.023

Gosselin, F. & Schyns, P. G. (2001). Bubbles: A technique to reveal the use of information in recognition. *Vision Research, 41*, 2261-2271. https://doi.org/10.1016/S0042-6989(01)00097-9

Green, A. D., Tripp, D. A., Sullivan, M. J. L., & Davidson, M. (2009). The relationship between empathy and estimates of observed pain. *Pain Medicine*, *10*(2), 381-392. http://doi.org/10.1111/j.1526-4637.2009.00563.x

Grimshaw, G. M., Bulman-Fleming, M. B., & Ngo, C. (2004). A signal-detection analysis of sex differences in the perception of emotional faces. *Brain and cognition*, *54*(3), 248-250. http://doi.org/10.1016/j.bandc.2004.02.029

Hadjistavropoulos, T., Craig, K. D., Duck, S., Cano, A., Goubert, L., Jackson, P. L., ... & Fitzgerald, T. D. (2011). A biopsychosocial formulation of pain communication. *Psychological bulletin*, *137*(6), 910. http://doi.org/10.1037/a0023876

Hadjistavropoulos, H. D., Craig, K. D., Grunau, R. E., & Whitfield, M. F. (1997). Judging pain in infants: behavioural, contextual, and developmental determinants. *Pain*, *73*(3), 319-324. http://doi.org/10.1016/s0304-3959(97)00113-9

Hall, J. A. (1978). Gender effects in decoding nonverbal cues. *Psychological bulletin*, *85*(4), 845. https://doi.org/10.1037/0033-2909.85.4.845

Hall, J. K., Hutton, S. B., & Morgan, M. J. (2010). Sex differences in scanning faces: Does attention to the eyes explain female superiority in facial expression recognition?. *Cognition & Emotion*, *24*(4), 629-637. http://doi.org/10.1080/02699930902906882

Handa, R. J., & McGivern, R. F. (2015). Steroid hormones, receptors, and perceptual and cognitive sex differences in the visual system. *Current Eye Research*, *40*(2), 110-127. http://doi.org/10.3109/02713683.2014.952826

Heisz, J. J., Pottruff, M. M., & Shore, D. I. (2013). Females scan more than males: a potential mechanism for sex differences in recognition memory. *Psychological science*, *24*(7), 1157-1163. http://doi.org/10.1177/0956797612468281

Hess, U., & Blairy, S. (2001). Facial mimicry and emotional contagion to dynamic emotional facial expressions and their influence on decoding accuracy. *International journal of psychophysiology*, *40*(2), 129-141. http://doi.org/10.1016/s0167-8760(00)00161-6

Hill, M. L., & Craig, K. D. (2004). Detecting deception in facial expressions of pain: accuracy and training. *The Clinical journal of pain*, *20*(6), 415-422. http://doi.org/10.1097/00002508-200411000-00006

Hirsh, A. T., George, S. Z., & Robinson, M. E. (2009). Pain assessment and treatment disparities: a virtual human technology investigation. *Pain*, *143*(1-2), 106-113. http://doi.org/10.1016/j.pain.2009.02.005

Jones, W., & Klin, A. (2013). Attention to eyes is present but in decline in 2–6-month-old infants later diagnosed with autism. *Nature*, *504*(7480), 427-431. http://doi.org/10.1038/nature12715
Jonides, J., & Irwin, D. E. (1981). Capturing attention. *Cognition*, *10*(1-3), 145-150. https://doi.org/10.1016/0010-0277(81)90038-X
Kang, J., Ham, B. J., & Wallraven, C. (2017). Cannot avert the eyes: Reduced attentional blink toward others' emotional expressions in empathic people. *Psychonomic bulletin & review*, *24*(3), 810-820. https://doi.org/10.3758/s13423-016-1171-x
Kayyal, M. H., & Russell, J. A. (2013). Americans and Palestinians judge spontaneous facial expressions of emotion. *Emotion*, *13*(5), 891. https://doi.org/10.1037/a0033244
Keogh, E. (2014). Gender differences in the nonverbal communication of pain: A new direction for sex, gender, and pain research?. *Pain*, 155(10), 1927-1931. https://doi.org/10.1016/j.pain.2014.06.024
Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3?.
Konrath, S. H., O'Brien, E. H., & Hsing, C. (2011). Changes in dispositional empathy in American college students over time: A meta-analysis. *Personality and Social Psychology Review*, *15*(2), 180-198. https://doi.org/10.1177/1088868310377395
Kret, M. E., & De Gelder, B. (2012). A review on sex differences in processing emotional signals. *Neuropsychologia*, *50*(7), 1211-1221. https://doi.org/10.1016/j.neuropsychologia.2011.12.022
Krill, A. L., Platek, S. M., & Wathne, K. (2008). Feelings of control during social exclusion are partly accounted for by empathizing personality. *Personality and Individual Differences*, *45*(7), 684–688. https://doi.org/10.1016/j.paid.2008.07.016
Krumhuber, E. G., Tamarit, L., Roesch, E. B., & Scherer, K. R. (2012). FACSGen 2.0 animation software: generating three-dimensional FACS-valid facial expressions for emotion research. *Emotion*, *12*(2), 351. https://doi.org/10.1037/a0026632
Kunz, M., & Lautenbacher, S. (2014). The faces of pain: a cluster analysis of individual differences in facial activity patterns of pain. European Journal of Pain, 18(6), 813-823. https://doi.org/10.1002/j.1532-2149.2013.00421.x
Kunz, M., Meixner, D., & Lautenbacher, S. (2019). Facial muscle movements encoding pain—a systematic review. Pain, 160(3), 535-549. https://doi.org/10.1097/j.pain.0000000000001424
Langner, O., Becker, E. S., & Rinck, M. (2009). Social anxiety and anger identification: Bubbles reveal differential use of facial information with low spatial frequencies. *Psychological science*, *20*(6), 666-670. https://doi.org/10.1111/j.1467-9280.2009.02357.x
Lawrence, E. J., Shaw, P., Baker, D., Baron-Cohen, S., & David, A. S. (2004). Measuring empathy: reliability and validity of the Empathy Quotient. Psychological medicine, 34(5), 911-920. https://doi.org/10.1017/S0033291703001624
Lee, J., Zaki, J., Harvey, P. O., Ochsner, K., & Green, M. F. (2011). Schizophrenia patients are impaired in empathic accuracy. *Psychological medicine*, *41*(11), 2297-2304.https://doi.org/10.1017/s0033291711000614
Lin, J., Drain, A., Goharzad, A., & Mende-Siedlecki, P. (2020). What factors fuel racial bias in pain perception and treatment?: A meta-analysis across 40 experimental studies.
Little, A. C. (2013). The influence of steroid sex hormones on the cognitive and emotional processing of visual stimuli in humans. *Frontiers in neuroendocrinology*, *34*(4), 315-328. https://doi.org/10.1016/j.yfrne.2013.07.009

Lucas-Molina, B., Pérez-Albéniz, A., Ortuño-Sierra, J., & Fonseca-Pedrero, E. (2017). Dimensional structure and measurement invariance of the Interpersonal Reactivity Index (IRI) across gender. *Psicothema*, *29*(4), 590-595. https://doi.org/10.7334/psicothema2017.19

Lutchmaya, S., Baron-Cohen, S., & Raggatt, P. (2002). Foetal testosterone and eye contact in 12-month-old human infants. *Infant Behavior and Development*, *25*(3), 327-335. https://doi.org/10.1016/S0163-6383(02)00094-2

Mandal, M. K., & Palchoudhury, S. (1985). Perceptual skill in decoding facial affect. *Perceptual and Motor Skills*, *60*(1), 96-98. https://doi.org/10.2466/pms.1985.60.1.96

McClure, E. B. (2000). A meta-analytic review of sex differences in facial expression processing and their development in infants, children, and adolescents. *Psychological bulletin*, *126*(3), 424. https://doi.org/10.1037/0033-2909.126.3.424

Meister, E., Horn-Hofmann, C., Kunz, M., Krumhuber, E. G., & Lautenbacher, S. (2021). Decoding of facial expressions of pain in avatars: does sex matter?. *Scandinavian Journal of Pain*, *21*(1), 174-182. https://doi.org/10.1515/sjpain-2020-0078

Mestre, M. V., Samper, P., Frías, M. D., & Tur, A. M. (2009). Are women more empathetic than men? A longitudinal study in adolescence. *The Spanish journal of psychology*, *12*(1), 76-83. https://doi.org/10.1017/s1138741600001499

Miron-Shatz, T., Ormianer, M., Rabinowitz, J., Hanoch, Y., & Tsafrir, A. (2020). Physician experience is associated with greater underestimation of patient pain. *Patient education and counseling*, *103*(2), 405-409. https://doi.org/10.1016/j.pec.2019.08.040

Montagne, B., Kessels, R. P., Frigerio, E., de Haan, E. H., & Perrett, D. I. (2005). Sex differences in the perception of affective facial expressions: Do men really lack emotional sensitivity?. *Cognitive processing*, *6*(2), 136-141. https://doi.org/10.1007/s10339-005-0050-6

Motley, M. T., & Camden, C. T. (1988). Facial expression of emotion: A comparison of posed expressions versus spontaneous expressions in an interpersonal communication setting. *Western Journal of Communication (includes Communication Reports)*, *52*(1), 1-22. https://doi.org/10.1080/10570318809389622

Naab, P. J., & Russell, J. A. (2007). Judgments of emotion from spontaneous facial expressions of New Guineans. *Emotion*, *7*(4), 736. https://doi.org/10.1037/1528-3542.7.4.736

Newman, T. M., Macomber, D., Naples, A. J., Babitz, T., Volkmar, F., & Grigorenko, E. L. (2007). Hyperlexia in children with autism spectrum disorders. *Journal of autism and developmental disorders*, *37*(4), 760-774. https://doi.org/10.1007/s10803-006-0206-y

Nopoulos, P., Flaum, M., O'Leary, D., & Andreasen, N. C. (2000). Sexual dimorphism in the human brain: evaluation of tissue volume, tissue composition and surface anatomy using magnetic resonance imaging. *Psychiatry Research: Neuroimaging*, *98*(1), 1-13. https://doi.org/10.1016/s0925-4927(99)00044-x

Palermo, R., & Coltheart, M. (2004). Photographs of facial expression: Accuracy, response times, and ratings of intensity. *Behavior Research Methods, Instruments, & Computers*, *36*(4), 634-638. https://doi.org/10.3758/bf03206544

Pelli, D. G., & Vision, S. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial vision*, *10*, 437-442. https://doi.org/10.1163/156856897x00366

Peterson, M. F., & Eckstein, M. P. (2012). Looking just below the eyes is optimal across face recognition tasks. *Proceedings of the National Academy of Sciences*, *109*(48), E3314-E3323. https://doi.org/10.1073/pnas.1214269109

Phillips, W. A., Chapman, K. L., & Berry, P. D. (2004). Size perception is less context-sensitive in males. *Perception*, *33*(1), 79-86. https://doi.org/10.1068/p5110

Plouffe-Demers, M. P., Fiset, D., Saumure, C., Duncan, J., & Blais, C. (2019). Strategy shift toward lower spatial frequencies in the recognition of dynamic facial expressions of basic emotions: when it moves it is different. *Frontiers in psychology*, *10*, 1563. https://doi.org/10.3389/fpsyg.2019.01563

Plouffe-Demers, M.-P., & Blais, C. (2022, June 16). Facial Expression of Pain: Sex Differences in the Discrimination of Varying Intensities. Retrieved from osf.io/9t6eu

Posner, M. I. (1980). Orienting of attention. *Quarterly journal of experimental psychology*, *32*(1), 3-25. https://doi.org/10.1080/00335558008248231

Prkachin, K. M. (1992). The consistency of facial expressions of pain: a comparison across modalities. Pain, 51(3), 297-306. https://doi.org/10.1016/0304-3959(92)90213-U

Prkachin, K. M., Currie, N. A., & Craig, K. D. (1983). Judging nonverbal expressions of pain. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, *15*(4), 409. https://doi.org/10.1037/h0080757

Prkachin, K. M., Mass, H., & Mercer, S. R. (2004). Effects of exposure on perception of pain expression. *Pain*, *111*(1-2), 8-12. https://doi.org/10.1016/j.pain.2004.03.027

Prkachin, K. M., & Solomon, P. E. (2008). The structure, reliability and validity of pain expression: Evidence from patients with shoulder pain. Pain, 139(2), 267-274. https://doi.org/10.1016/j.pain.2008.04.010

Pronina, I., & Rule, N. O. (2014). Inducing bias modulates sensitivity to nonverbal cues of others' pain. *European Journal of Pain*, *18*(10), 1452-1457. https://doi.org/10.1002/ejp.510

Reniers, R. L., Corcoran, R., Drake, R., Shryane, N. M., & Völlm, B. A. (2011). The QCAE: A questionnaire of cognitive and affective empathy. *Journal of personality assessment*, *93*(1), 84-95. https://doi.org/10.1080/00223891.2010.528484

Richoz, A. R., Jack, R. E., Garrod, O. G., Schyns, P. G., & Caldara, R. (2015). Reconstructing dynamic mental models of facial expressions in prosopagnosia reveals distinct representations for identity and expression. *Cortex*, *65*, 50-64. https://doi.org/10.1016/j.cortex.2014.11.015

Riva, P., Sacchi, S., Montali, L., & Frigerio, A. (2011). Gender effects in pain detection: Speed and accuracy in decoding female and male pain expressions. *European Journal of Pain*, *15*(9), 985-e1. https://doi.org/10.1016/j.ejpain.2011.02.006

Robinson, M. E., & Wise, E. A. (2003). Gender bias in the observation of experimental pain. *Pain*, *104*(1-2), 259-264. https://doi.org/10.1016/s0304-3959(03)00014-9

Robinson, K., Blais, C., Duncan, J., Forget, H., & Fiset, D. (2014). The dual nature of the human face: there is a little Jekyll and a little Hyde in all of us. *Frontiers in psychology*, *5*, 139. https://doi.org/10.3389/fpsyg.2014.00139

Rogers, S. L., Speelman, C. P., Guidetti, O., & Longmuir, M. (2018). Using dual eye tracking to uncover personal gaze patterns during social interaction. *Scientific reports*, *8*(1), 1-9. https://doi.org/10.1038/s41598-018-22726-7

Royer, J., Blais, C., Charbonneau, I., Déry, K., Tardif, J., Duchaine, B., ... & Fiset, D. (2018). Greater reliance on the eye region predicts better face recognition ability. *Cognition*, *181*, 12-20. https://doi.org/10.1016/j.cognition.2018.08.004

Royer, J., Blais, C., Gosselin, F., Duncan, J., & Fiset, D. (2015). When less is more: Impact of face processing ability on recognition of visually degraded faces. J Exp Psychol Hum Percept Perform, 41(5), 1179. https://doi.org/10.1037/xhp0000095

Rueckert, L., Branch, B., & Doan, T. (2011). Are gender differences in empathy due to differences in emotional reactivity?. *Psychology*, *2*(6), 574. https://doi.org/10.4236/psych.2011.26088

Saarela, M. V., Hlushchuk, Y., Williams, A. C. D. C., Schürmann, M., Kalso, E., & Hari, R. (2007). The compassionate brain: humans detect intensity of pain from another's face. *Cerebral cortex*, *17*(1), 230-237. https://doi.org/10.1093/cercor/bhj141

Sæther, L., Van Belle, W., Laeng, B., Brennen, T., & Øvervoll, M. (2009). Anchoring gaze when categorizing faces' sex: evidence from eye-tracking data. *Vision research*, *49*(23), 2870-2880.https://doi.org/10.1016/j.visres.2009.09.001

Samson, A. C. (2012). The influence of empathizing and systemizing on humor processing: Theory of mind and humor. *Humor, 25(1)*. https://doi.org/10.1515/humor-2012-0005

Saumure, C., Plouffe-Demers, M. P., Estéphan, A., Fiset, D., & Blais, C. (2018). The use of visual information in the recognition of posed and spontaneous facial expressions. *Journal of vision*, *18*(9), 21-21. https://doi.org/10.1167/18.9.21

Schwarzkopf, D. S., Song, C., & Rees, G. (2011). The surface area of human V1 predicts the subjective experience of object size. *Nature neuroscience*, *14*(1), 28-30.https://doi.org/10.1038/nn.2706

Shaqiri, A., Roinishvili, M., Grzeczkowski, L., Chkonia, E., Pilz, K., Mohr, C., ... & Herzog, M. H. (2018). Sex-related differences in vision are heterogeneous. *Scientific reports*, *8*(1), 1-10.https://doi.org/10.1038/s41598-018-25298-8

Simon, D., Craig, K. D., Gosselin, F., Belin, P., & Rainville, P. (2008). Recognition and discrimination of prototypical dynamic expressions of pain and emotions. *PAIN®*, *135*(1-2), 55-64.https://doi.org/10.1016/j.pain.2007.05.008

Simon, D., Craig, K. D., Miltner, W. H., & Rainville, P. (2006). Brain responses to dynamic facial expressions of pain. *Pain*, *126*(1-3), 309-318.https://doi.org/10.1016/j.pain.2006.08.033

Simoncelli, E. P. (1999). Image and multi-scale pyramid tools [Computer software]. *NewYork.*

Smith, F. G., Jones, B. C., & DeBruine, L. M. (2010). Individual differences in empathizing and systemizing predict variation in face preferences. *Personality and Individual Differences*, *49*(6), 655–658. https://doi.org/10.1016/j.paid.2010.05.023

Singer, T., Seymour, B., O'doherty, J., Kaube, H., Dolan, R. J., & Frith, C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science*, *303*(5661), 1157-1162. http://doi.org/10.1126/science.1093535

Sokhn, N., Bertoli, F., & Caldara, R. (2017, February). A left eye bias for female faces. In *2017 9th International Conference on Knowledge and Smart Technology (KST)* (pp. 274-278). IEEE. https://doi.org/10.1109/kst.2017.7886129

Song, C., Schwarzkopf, D. S., & Rees, G. (2013). Variability in visual cortex size reflects tradeoff between local orientation sensitivity and global orientation modulation. *Nature communications*, *4*(1), 1-10. https://doi.org/10.1038/ncomms3201

Spezio, M. L., Adolphs, R., Hurley, R. S., & Piven, J. (2007). Abnormal use of facial information in high-functioning autism. *Journal of autism and developmental disorders*, *37*(5), 929-939. https://doi.org/10.1007/s10803-006-0232-9

Tardif, J., Fiset, D., Zhang, Y., Estéphan, A., Cai, Q., Luo, C., ... & Blais, C. (2017). Culture shapes spatial frequency tuning for face identification. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(2), 294.https://doi.org/10.1037/xhp0000288

Tessier, M. H., Gingras, C., Robitaille, N., & Jackson, P. L. (2019). Toward dynamic pain expressions in avatars: perceived realism and pain level of different action unit orders. *Computers in Human Behavior*, *96*, 95-109.https://doi.org/10.1016/j.chb.2019.02.001

Thayer, J., & Johnsen, B. H. (2000). Sex differences in judgement of facial affect: A multivariate analysis of recognition errors. *Scandinavian journal of psychology*, *41*(3), 243-246.https://doi.org/10.1111/1467-9450.00193

Thompson, A. E., & Voyer, D. (2014). Sex differences in the ability to recognise non-verbal displays of emotion: A meta-analysis. *Cognition and Emotion*, *28*(7), 1164-1195.https://doi.org/10.1080/02699931.2013.875889

Tingley D, Yamamoto T, Hirose K, Keele L, Imai K (2013). mediation: R Package for Causal Mediation Analysis. R package version 4.4.2, URL http://CRAN.R-project.org/ package=mediation.

Van der Graaff, J., Branje, S., De Wied, M., Hawk, S., Van Lier, P., & Meeus, W. (2014). Perspective taking and empathic concern in adolescence: gender differences in developmental changes. *Developmental psychology*, *50*(3), 881. https://doi.org/10.1037/a0034325

Vanston, J. E., & Strother, L. (2017). Sex differences in the human visual system. *Journal of neuroscience research*, *95*(1-2), 617-625. https://doi.org/10.1002/jnr.23895

Vassallo, S., Cooper, S. L., & Douglas, J. M. (2009). Visual scanning in the recognition of facial affect: Is there an observer sex difference?. *Journal of Vision*, *9*(3), 11-11.https://doi.org/10.1167/9.3.11

Wagner, H. L., MacDonald, C. J., & Manstead, A. S. (1986). Communication of individual emotions by spontaneous facial expressions. *Journal of personality and social psychology*, *50*(4), 737.https://doi.org/10.1037/0022-3514.50.4.737

Wandner, L. D., Stutts, L. A., Alqudah, A. F., Craggs, J. G., Scipio, C. D., Hirsh, A. T., & Robinson, M. E. (2010). Virtual human technology: patient demographics and healthcare training factors in pain observation and treatment recommendations. *Journal of pain research*, *3*, 241.https://doi.org/10.2147/jpr.s14708

Watson, A. B., & Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Perception & psychophysics*, *33*(2), 113-120.https://doi.org/10.3758/bf03202828

Whittle, S., Yücel, M., Yap, M. B., & Allen, N. B. (2011). Sex differences in the neural correlates of emotion: evidence from neuroimaging. *Biological psychology*, *87*(3), 319-333.https://doi.org/10.1016/j.biopsycho.2011.05.003

Williams, A. C. D. C. (2002). Facial expression of pain: an evolutionary account. *Behavioral and brain sciences*, *25*(4), 439-455.https://doi.org/10.1017/s0140525x02000080

Willenbockel, V., Sadr, J., Fiset, D., Horne, G. O., Gosselin, F., & Tanaka, J. W. (2010). Controlling low-level image properties: the SHINE toolbox. *Behavior research methods*, *42*(3), 671-684.https://doi.org/10.3758/brm.42.3.671

Wingenbach, T. S., Ashwin, C., & Brosnan, M. (2018). Sex differences in facial emotion recognition across varying expression intensity levels from videos. *PLoS one*, *13*(1), e0190634. https://doi.org/10.1371/journal.pone.0190634

Wells, J. C. (2007). Sexual dimorphism of body composition. *Best practice & research Clinical endocrinology & metabolism*, *21*(3), 415-430.https://doi.org/10.1016/j.beem.2007.04.007

Wood, W., & Eagly, A. H. (2012). Biosocial construction of sex differences and similarities in behavior. In Advances in experimental social psychology (Vol. 46, pp. 55-123). Academic Press. https://doi.org/10.1016/B978-0-12-394281-4.00002-7

Yamada, M., & Decety, J. (2009). Unconscious affective processing and empathy: an investigation of subliminal priming on the detection of painful facial expressions. *Pain*, *143*(1-2), 71-75. https://doi.org/10.1016/j.pain.2009.01.028

Yitzhak, N., Pertzov, Y., & Aviezer, H. (2021). The elusive link between eye-movement patterns and facial expression recognition. *Social and Personality Psychology Compass*, e12621.https://doi.org/10.1111/spc3.12621